

## 空間スケーラビリティを用いた画像解像度変換法

石川 孝明<sup>†</sup> 渡辺 裕<sup>†</sup>(正会員)<sup>†</sup> 早稲田大学大学院国際情報通信研究科

あらまし 近年、画像符号化におけるスケーラビリティの重要性が高まっている。特に、空間スケーラビリティを有する JPEG 2000 は、デジタルシネマで使用される符号化方式として採用され、今後の普及が期待されている。しかし、JPEG 2000 の空間スケーラビリティは、オクターブ分割を用いるサブバンド符号化により実現しているため、低域画像の解像度が  $1/2^n$  に制限される。そのため、現行の表示デバイスでデジタルシネマのコンテンツを視聴するには、画素アスペクト比の相違を考慮した有理数比の解像度変換を行う必要がある。そこで本論文では、JPEG 2000 の空間スケーラビリティにより得られる低域画像と、有理数比の縮小を可能とするデシメーションフィルタを組み合わせた解像度変換法を提案する。本提案手法を用いることで、デシメーションフィルタのみを用いる従来手法と比較し、1~2[dB] 程度の画質改善を実現した。

キーワード：解像度変換，空間スケーラビリティ，デジタルシネマ，JPEG 2000

**Summary** Recently, the importance of SNR and spatial scalability in image coding has been increased. JPEG 2000 has advanced SNR and spatial scalability functions. However, the spatial scalability of JPEG 2000 is limited to the cases where decimated resolutions are  $1/2^n$  of the source image resolution. Therefore, Digital-Cinema resolution cannot be converted to SDTV or QVGA size image by decoding the JPEG 2000 bitstream directly. We propose a resolution conversion method using spatial scalability, which combines LL sub-band with a decimation filter. Experimental results confirm an improvement in PSNR of 1-2 dB as compared to conventional resolution conversion methods.

**Key words:** Resolution conversion, Spatial scalability, Digital Cinema, JPEG 2000

## 1. はじめに

近年、マルチメディアコンテンツの再生環境が多様化したことを背景に、符号化されたデジタルコンテンツのシングルソース・マルチユースへの期待が高まっている。このようなアプリケーションを実現する技術として、ビデオトランスコーダ<sup>1)</sup>やサブバンド符号化によるスケーラビリティ<sup>2)</sup>などがある。

ISO/IEC JTC 1/SC 29/WG 1 (JPEG) および SMPTE (Society of Motion Picture and Television Engineers) において、現在標準化が進められているデジタルシネマは、フィルムを媒体とする従来の映画シ

テムではなく、超高解像度のデジタル画像を利用するシステムである。デジタルシネマでは、各劇場の再生環境に柔軟性を持たせる目的から、現状の技術で比較的容易に実現可能な 2K (2048 × 1080 画素の映像を上映可能) のプロジェクタと、フィルムとほぼ同じ画質を再現するとされる 4K (4096 × 2160 画素の映像を上映可能) のプロジェクタ、の二つの環境を想定している<sup>3)</sup>。これらの点から、デジタルシネマは、シングルソース・マルチユースを実現した実用的なアプリケーションであると言える。

デジタルシネマの符号化には、超高解像度の映像符号化と空間解像度へのスケーラビリティが必要とされている。これらを実現する符号化方式として、JPEG 2000 がある<sup>4)</sup>。JPEG 2000 は、2000 年に国際標準化され、2004 年にデジタルシネマの符号化方式として採用された。JPEG 2000 の空間スケーラビリティを利用するこ

“Image Resolution Conversion Method using Spatial Scalability” by Takaaki ISHIKAWA, and Hiroshi WATANABE, (Member), (Graduate School of Global Information and Telecommunication Studies, Waseda University).

とで、4096 × 2160 画素のプロジェクタと 2048 × 1080 画素のプロジェクタの双方に対応することが可能である。

一方で、JPEG 2000 のもつ空間スケーラビリティは、デコード後の画像解像度が  $1/2^n$  に制限される。これは、JPEG 2000 で利用されているウェーブレット変換とオクターブ分割によるサブバンド符号化に起因するもので、オクターブ分割を行う以上、避けることが困難である。デジタルシネマでは、想定される 2 種類の空間解像度が共に  $1/2^n$  の関係にあるため、問題にはならないが、現行の表示デバイスでの再生を考慮する場合には問題となる。そのため、シングルソース・マルチユースの適用範囲を劇場だけに限らず、ホームシアターやテレビ放送にまで広げ、デジタルシネマを携帯電話や SDTV, HDTV など視聴するためには、 $1/2^n$  の制限を受けない有理数比の解像度変換が必要となる。

この課題に対する解として、 $M : N$  の変換フィルタバンクを用いた、有理数比の解像度変換を実現する手法がある<sup>5),6)</sup>。しかし、コンテンツの再生環境に複数の空間解像度が存在することを考慮すれば、すべての空間解像度に対応する  $M$  を用いることは現実的ではない。これは、符号化において DCT を用いるシステム<sup>7)</sup>においても同様の問題である。

そこで本論文では、空間スケーラビリティにより得られる低域画像と、デシメーションフィルタを組み合わせた有理数比の解像度変換法を提案する。また、提案手法を用いたデジタルシネマを対象とした実装システムを提案する。提案手法は、分離型 2 次元システムでフィルタ処理を行うため、画素アスペクト比変換を容易に行うことができる。それと同時に、本手法により、原画像にデシメーションフィルタを直接施すよりも高画質の画像を得ることができる。

以下に本論文の構成を示す。2. において、空間スケーラビリティにより得られる低域画像と、デシメーションフィルタを組み合わせたマルチアウトプット解像度変換手法について述べる。3. では、提案する解像度変換手法を用い、画素アスペクト比の変換を考慮した実装システムについて述べる。4. では、評価実験により提案手法の有効性について述べ、最後に、5. において、本論文のまとめを述べる。

## 2. マルチアウトプット解像度変換手法

本章では、デジタルシネマの符号化方式として JPEG 2000 が利用されることを前提として、JPEG 2000 の空間スケーラビリティにより取得可能な低域画像とデシメーションフィルタを組み合わせた有理数比の解像度変換手法について述べる。

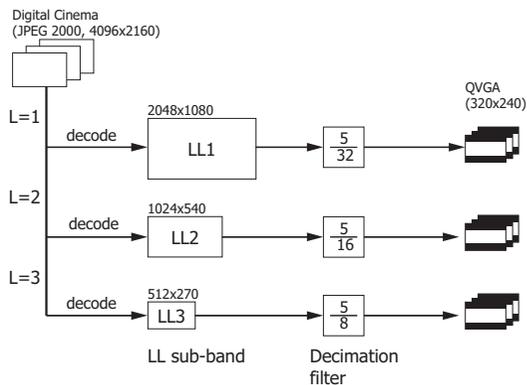


図 1 低域画像とフィルタの組み合わせ例

Fig. 1 Decoding path from 4096x2160 to 320x240

### 2.1 低域画像とデシメーションフィルタ

本提案手法における低域画像とフィルタの組み合わせは、任意の空間解像度を持つ映像を取得する場合に、複数の組み合わせが考えられる。これは、空間スケーラビリティにより取得可能な縮小画像の種類が、目的の縮小画像の空間解像度が低いほど、多様性を増すためである。

例えば、デジタルシネマサイズ (4096 × 2160 画素) から QVGA サイズ (320 × 240 画素) の縮小画像を得る場合、低域画像として取得可能な解像度は、図 1 に示したように 3 通り存在する。すなわち、選択可能な低域画像は、ウェーブレット分割の回数  $L$  が 1 から 3 回の場合である。これらの低域画像に対し、縮小率の異なるデシメーションフィルタを用いて縮小処理を行うことで、最終的に、QVGA サイズの映像を取得するための組み合わせが 3 通り定義される。 $L = 4$  の低域画像を利用した縮小処理は実現できないため、選択可能な分割数  $L$  は、入力画像の解像度と出力画像の解像度に制限される。

### 2.2 Lanczos フィルタ

本提案手法では、デシメーションフィルタとして Lanczos フィルタを用いる。Lanczos フィルタは、窓関数が 0 でない区間の長さにより、複数の種類が存在する。この区間のうち、正の領域の区間長をローブ幅とよび、本論文ではこれを  $N$  で表す。 $N = 3$  の場合を特に Lanczos3 (3-lobed Lanczos-windowed sinc function) フィルタと呼ぶ<sup>8)</sup>。Lanczos3 フィルタを用いた有理数比の縮小処理例を、縮小率が  $4/5$  の場合について図 2 に示す。

出力信号系列中における信号  $\Delta 2$  は、入力信号の 2 と 3 の中心に位置する離散信号に対応する。この中心位置に Lanczos3 フィルタを対応させ、Lanczos3 フィルタ関数上で入力信号に対応するの値を入力信号に重み付けし、 $\Delta 2$  の信号の値を算出する。

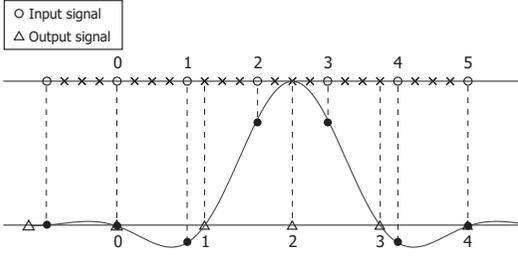


図2 Lanczos3 フィルタによる画素間引き処理  
Fig. 2 Decimation with Lanczos3 filter

以下に、Lanczos フィルタのデシメーションフィルタとしての特徴について述べる。

### 2.2.1 Lanczos フィルタの周波数特性

カットオフ周波数  $\omega_c$  を  $\pi/2$  とした Lanczos フィルタの周波数特性を、図3に示す。比較のために、JPEG 2000 で定義される 5/3-Integer フィルタと 9/7-Daubechies フィルタを同図に示す。図3に示されたように、縮小率が  $1/2$  の場合の Lanczos フィルタはカットオフ周波数付近での急峻性が最も高く、サイドローブにリップルが生じているが、他のフィルタと比較してより多くの減衰量を得ていることがわかる。また、ガウシアンフィルタなどのデシメーションフィルタに対する優位性もすでに明らかにされている<sup>8)</sup>。

### 2.2.2 カットオフ周波数とタップ数の関係

縮小率  $r$  の Lanczos フィルタのカットオフ周波数は、ロープ幅  $N$  の値によらず、 $\omega_c = \pi r$  で与えられる。したがって、任意のカットオフ周波数を与える Lanczos フィルタのタップ数は、式(1)で表される。

$$T_{Lanc}(N) = 2 \lfloor \frac{N}{r} \rfloor + 1 = 2 \lfloor \frac{\pi}{\omega_c} N \rfloor + 1 \quad (1)$$

式(1)は、Lanczos フィルタのタップ数が縮小率により決定されることを意味している。

一般に、フィルタのタップ数が増加するとフィルタリングに要する演算量も増加する。Lanczos フィルタをローパスフィルタとして用いた場合、フィルタのタップ数が、要求されるカットオフ周波数  $\omega_c$  (もしくは縮小率  $r$ ) に依存するため、縮小率が上がることでタップ数が増加し、それに伴い演算量も増加する。

### 2.2.3 理想ローパスフィルタと Lanczos フィルタの関係

Lanczos フィルタは、理想ローパスフィルタである sinc 関数に対し、式(2)の窓関数を掛け合わせた式(3)で定義される。実際のフィルタ係数は、式(3)で描かれる関数のうち、関数の値が0では無い  $-N < x < N$  の区間を、式(1)で算出されるタップ数により等分割する

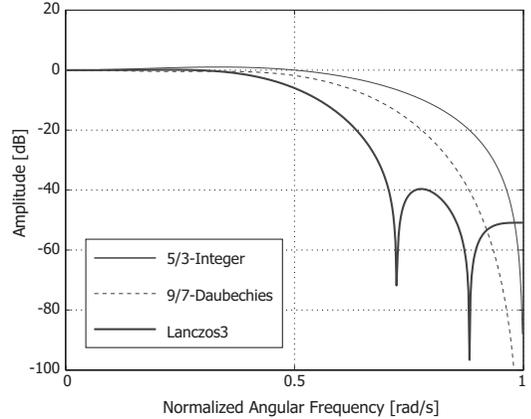


図3 Lanczos3 フィルタの周波数特性 (縮小率  $1/2$ )  
Fig. 3 Frequency response of Lanczos3 filter ( $r = 1/2$ )

ことで得られる。

$$W_{Lanc.}(x) = \begin{cases} \frac{\sin(\frac{\pi x}{N})}{\pi \frac{x}{N}}, & |x| < N \\ 0, & otherwise \end{cases} \quad (2)$$

$$LanczosN(x) = \begin{cases} \frac{\sin(\pi x)}{\pi x} \frac{\sin(\frac{\pi x}{N})}{\pi \frac{x}{N}}, & |x| < N \\ 0, & otherwise \end{cases} \quad (3)$$

ロープ幅  $N$  で表される式(3)は、 $N$  を無限大に近づければ sinc 関数と等しくなる。したがって、 $N$  の値が増加することで、理想的なローパスフィルタに近づく。しかしその一方で、フィルタのタップ数が増大し、演算量が無限に発散することも式(1)から明らかである。したがって、画質改善を目的として、デシメーションフィルタにロープ幅  $N$  を単純に増加させた Lanczos フィルタを利用することは現実的ではない。

### 2.3 フィルタリングにおける演算コストの定式化

ロープ幅  $N$  の Lanczos フィルタを、低域画像と組み合わせるデシメーションフィルタとして使用するためには、演算コストの面からの検討が不可欠であることを前節で明らかとした。本節では、ロープ幅  $N$  の Lanczos フィルタの演算量を、ポリフェーズ構成で最適化した回路の加算および乗算回数により定式化する。これにより、提案手法の解像度変換が、デシメーションフィルタのみを用いる場合と比較し、低い演算量の処理を保証できる。以下に、フィルタリングに用いる回路のポリフェーズ構成による最適化と、それをを用いた演算コストの定式化について述べる。

#### 2.3.1 ポリフェーズ構成による最適化

一般に、有理数比  $U/D$  の縮小処理を実現するデシメータは、分割数が  $U$  のポリフェーズ構成に分解すること

ができる<sup>9)</sup>。ポリフェーズ分解には、式(4)に示すポリフェーズ表現を用いる。

$$H(z) = \sum_{n=-\infty}^{\infty} h(n)z^{-n} = \sum_{i=0}^{U-1} R_i(z^U)z^{-(U-i-1)} \quad (4)$$

ただし、

$$R_i(z) = \sum_{n=-\infty}^{\infty} h(Un - i - 1)z^{-n}$$

式(4)のポリフェーズ構成は図4-(a)のような回路図となる。有理数比  $U/D$  に  $tD-1 = kU$  なる関係が成立すると仮定し、マルチレートシステムの等価関係を用いると、図4-(a)は、図4-(b)、(c)と変形でき、最終的に図4-(d)の構成で表すことができる。図4-(d)における各チャンネルのポリフェーズフィルタ  $R_i(z)$  ( $i \in 1, 2, \dots, U-1$ ) とダウンサンプラの組み合わせは、直接構成の FIR フィルタに置き換えることが可能である。そのため、乗算器の個数を減少させることができる。ここで、 $t = 1$  とすれば、各チャンネルのアップサンプラと遅延器は、コミュテータ (commutator)<sup>9)</sup>により代用できることが容易にわかる。これにより大幅な処理の削減を行える。また、ソフトウェアにおいては、各チャンネルを独立に設計できるため並列処理が可能となる。

Lanczos フィルタは、有理数比  $U/D$  の縮小処理を実現するデシメータであるため、その最小構成はポリフェーズ構成により与えられる。Lanczos フィルタの伝達関数  $H(z)$  は、カットオフ周波数が  $\omega_c = \pi/D$  を満たす。また、この伝達関数  $H(z)$  を  $U$  個にポリフェーズ分解したフィルタにより実際のフィルタリングを行う。

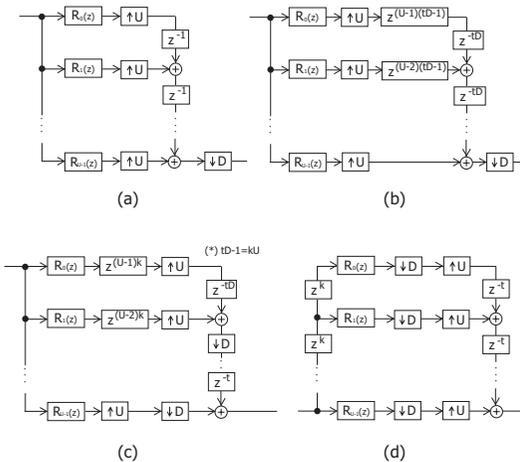


図4 デシメータのポリフェーズ構成

Fig. 4 Polyphase structure of multirate system

### 2.3.2 演算回数の定式化

有理数比  $U/D$  の解像度変換を実現する、ポリフェーズ構成により最適化されたデシメーションフィルタは、縮小率の分子の値  $U$  により回路構成に次のような違いがあるため、場合分けを行う。

・  $U$  が偶数の場合

ポリフェーズフィルタのチャンネルが偶数個存在し、 $phase = 0$  と  $\pi$  に対応するチャンネルでは対称性を利用する。 $phase = 0$  では奇数タップフィルタの対称性、 $phase = \pi$  では、偶数タップフィルタの対称性を利用する。

・  $U$  が奇数の場合

ポリフェーズフィルタのチャンネルが奇数個存在し、 $phase = 0$  に対応するチャンネルで、奇数タップフィルタの対称性を利用する。

また、ポリフェーズフィルタの各チャンネルからの出力を統合する加算器は、アップサンプラと遅延器をコミュテータで代用することで省略する。

さらに、入力信号数  $Src$  に対し、出力信号数  $Dst$  を出力するためにそれぞれのチャンネルが動作する合計回数は、式(5)の関係からチャンネルごとに  $Src$  と  $Dst$  の最大公約数である  $\gcd(Src, Dst)$  回である。また、 $Dst$  が  $U$  で割り切れることから、全入力信号に対する各チャンネルの総動作回数は等しい。

$$r = \frac{Dst}{Src} = \frac{U}{D} \cdot \frac{\gcd(Src, Dst)}{\gcd(Src, Dst)} = \frac{Dst}{U} = \gcd(Src, Dst) \quad (5)$$

以上のような性質を利用することにより、Lanczos フィルタの演算コストを、ロープ幅  $N$ 、出力信号の水平画素数  $X$ 、出力信号の垂直画素数  $Y$ 、水平方向の縮小率  $U_H/D_H$ 、垂直方向の縮小率  $U_V/D_V$  を用いて、式(6)~式(9)のように定式化する。

・  $U$  が偶数の場合

$$Cost_{multi} = XY(2(\lfloor \frac{D_H}{U_H} N \rfloor + 1)(1 - \frac{1}{U_H}) + 2(\lfloor \frac{D_V}{U_V} N \rfloor + 1)(1 - \frac{1}{U_V})) \quad (6)$$

$$Cost_{add} = XY((2\lfloor \frac{D_H}{U_H} N \rfloor + 1) - \frac{1}{U_H} + (2\lfloor \frac{D_V}{U_V} N \rfloor + 1) - \frac{1}{U_V}) \quad (7)$$

・  $U$  が奇数の場合

$$Cost_{multi} = XY((\lfloor \frac{D_H}{U_H} N \rfloor + 1)(2 - \frac{1}{U_H}) + (\lfloor \frac{D_V}{U_V} N \rfloor + 1)(2 - \frac{1}{U_V})) \quad (8)$$

$$Cost_{add} = XY \left( \left( 2 \left\lfloor \frac{D_H}{U_H} N \right\rfloor + 1 \right) - \frac{1}{U_H} \right) + \left( 2 \left\lfloor \frac{D_V}{U_V} N \right\rfloor + 1 \right) - \frac{1}{U_V} \quad (9)$$

2.4 最大分割数とデシメーションフィルタの縮小率  
提案手法における, *Lanczos* フィルタに入力される画素数は, ウェーブレット変換の分割数  $L$  を用いて式 (10) で表される. したがって, *Lanczos* フィルタの縮小率  $r_{Lanczos}$  と全体の縮小率  $r_{system}$  には, 式 (11) の関係が成立する. また, *Lanczos* フィルタの縮小率  $r_{Lanczos}$  は,  $r_{Lanczos} < 1$  を満たす. そのため, 提案手法で選択可能な入力画像の低域分割数  $L$  は, 式 (12) を満たさなければならない.

$$r_{Lanczos} = \frac{\text{Lanczos フィルタへの入力画素数}}{\text{システムへの入力画素数}} = \frac{\quad}{2^L} \quad (10)$$

$$r_{Lanczos} = \frac{\text{システムの出力画素数}}{\text{Lanczos フィルタへの入力画素数}} = \frac{\text{システムの出力画素数}}{\text{システムへの入力画素数}} \times 2^L = r_{system} \times 2^L \quad (11)$$

$$L < -\log_2(r_{system}) \quad (12)$$

### 2.5 演算量の拘束条件

本節では, 2.3 節で定式化した *Lanczos* フィルタの演算コストを利用し, 解像度変換処理全体の演算量について, *Lanczos* フィルタのローブ幅  $N$  に対する拘束条件を導出する. まず, 式 (6) と (8) で表される *Lanczos* フィルタの演算コストは, 式 (5) および式 (11) を用いることで, システムの入出力画素数である  $Src(H), Dst(H)$  および  $Src(V), Dst(V)$  とローブ幅  $N$ , ウェーブレット変換の分割数  $L$  を用いて, 式 (13) および式 (14) のように表現される.

$$Cost_{system}(L, N) = Dst(H)Dst(V)(Cost_{sub}(H, N) + Cost_{sub}(V, N)) \quad (13)$$

$$Cost_{sub}(p, N) = \begin{cases} 2 \left( \left\lfloor \frac{Src(p)}{Dst(p)} \frac{N}{2^L} \right\rfloor + 1 \right) \left( 1 - \frac{1}{U} \right), & U = \text{Even Number} \\ \left( \left\lfloor \frac{Src(p)}{Dst(p)} \frac{N}{2^L} \right\rfloor + 1 \right) \left( 2 - \frac{1}{U} \right), & U = \text{Odd Number} \end{cases}$$

ただし,  $U = \frac{Dst(p)}{\text{gcd}(Src(p)/2^L, Dst(p))}$  (14)

ここで, 分割数として  $L_a < L_b$  を満たす分割数  $L_a$  と分割数  $L_b$  が存在すると仮定する. 分割数  $L_a$  の場合には, 使用する *Lanczos* フィルタのローブ幅を  $N$  とし, 分割数  $L_b$  の場合には  $N + \alpha$  とする.  $\alpha$  は正の整数であり, ローブ幅の増加数を表す変数である. 分割数  $L_a$  と分割数  $L_b$  の双方における演算量を比較すると, 式 (15)

表 1  $\alpha$  の最大値

Table 1 Maximum value of alpha

$L_a \setminus L_b$	1	2	3	4
0	2	8	21	43
1		2	9	20
2			3	8
3				2

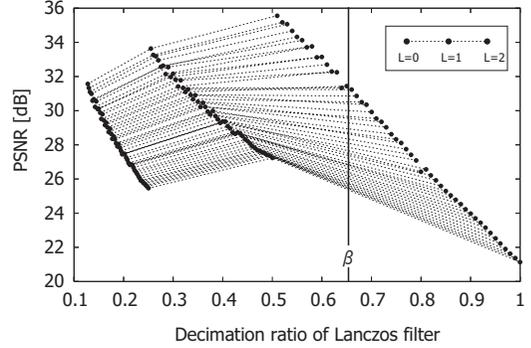


図 5  $r_{Lanczos}$  に対する品質変化 ( $L_{max} = 2$ )  
Fig. 5 Shift of PSNR due to decimation ratio of Lanczos filter ( $L_{max} = 2$ )

に示される不等式を満たす  $\alpha$  が存在するならば, 分割数が  $L_b$  の場合に,  $L_a$  の場合と比較して低い演算量の処理が行える. これと同時に *Lanczos* フィルタのローブ幅が増加し, 出力画像の品質が向上する.

$$Cost_{system}(L_a, N) > Cost_{system}(L_b, N + \alpha) \quad (15)$$

表 1 に  $\alpha$  の具体例を示す. 表 1 は,  $4096 \times 2160$  画素のデジタルシネマサイズからレターボックス方式で QCIF サイズ ( $176 \times 100$  画素) への解像度変換における最大ローブ増加数  $\alpha$  の値である. ただし, 初期値として  $N = 3$  を用いる. 表 1 より, 分割数  $L$  が増加するならばより大きいローブ幅の *Lanczos* フィルタを利用できることがわかる.

### 2.6 縮小率の拘束条件

本提案手法は, *Lanczos* フィルタのローブ幅  $N$  が大きければ, 解像度変換後の画像がより高画質となる. 一方で, ウェーブレット変換の分割数  $L$  が大きくなるにつれ, 解像度変換後の画像の品質が劣化する. これは, ダウンサンプリングによるエイリアシングが, 分割数  $L$  の増加と共に増すためである. したがって, 必ずしも選択可能な最大の分割数  $L_{max}$  とローブ幅  $N + \alpha$  の *Lanczos* フィルタとの組み合わせが最良であるとは限らない. これを定量的に示すために, 入力画像が  $1600 \times 1600$  画素のゾーンプレートを用い, 出力画像を  $60 \times 60$  画素から  $800 \times 800$  画素まで変化させた場合の出力画像の品質変化を測定した. 図 5 に,  $L_{max} = 2$  における実験

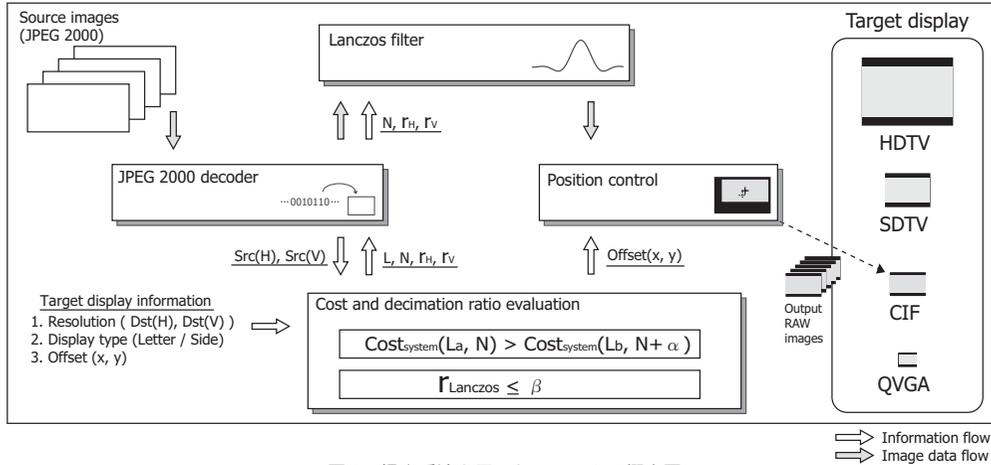


図 6 提案手法を用いたシステムの概念図  
Fig. 6 Proposed resolution conversion system

表 2  $\beta$  の値

Table 2 Values of beta

$L_{max}$	1	2	3	4
$\beta$	0.95	0.65	0.55	0.5

結果を示す．図 5 の横軸は Lanczos フィルタの縮小率  $r_{Lanczos}$  を示している．出力解像度が同一であるプロット点は点線で結ばれており， $L = 0$  の場合がグラフにおいて左側に位置し， $L = 2$  のプロット点は右側に位置している．図 5 より，出力解像度が同一であれば， $L = 0$  よりも  $L = 1$  がより高品質であるが， $L = 1$  と  $L = 2$  の比較では， $\beta$  を境界として品質が劣化していることがわかる．このように Lanczos フィルタの縮小率に着目することで， $\beta \leq r_{Lanczos} < 1$  の範囲において， $L = 2$  よりも  $L = 1$  を用いる組み合わせがより高画質の解像度変換を実現することがわかる．そこで，本提案手法では， $r_{Lanczos}$  の取りうる値として，式 (16) を拘束条件として定義する．ただし， $\beta$  は， $L_{max}$  の値により異なる値を持つ．表 2 に，ゾーンプレートに対する実験による  $\beta$  の値を示す．

$$r_{Lanczos} \leq \beta \quad (16)$$

### 3. 画素アスペクト比を考慮した実装システム

本章では，第 2 章で示したマルチアウトプット解像度変換法を用いた画素アスペクト比変換を含めた実装システムについて述べる．

#### 3.1 実装システムの概要

本提案システムは，ウェーブレット変換と縮小処理が分離型 2 次元システムにより構成されることを利用し，縦方向と横方向で縮小率を変え，画素アスペクト比の変換を正しく行う．図 6 に本提案システムの概要図を示す．

以下に，一連の処理過程について述べる．

**手順 1** 入力画像の水平および垂直解像度を取得し，出力画像の水平および垂直解像度，表示方式の種類を演算コスト評価部に入力する．

**手順 2** 演算コスト評価部では，組み合わせ可能な低域画像と Lanczos フィルタの組み合わせを，演算コストと Lanczos フィルタの縮小率の観点から評価し，低域画像の分割数  $L$  と Lanczos フィルタのローブ幅  $N$  を決定する．

**手順 3** JPEG 2000 デコーダにおいて，分割数  $L$  の低域成分を取得する．取得した低域画像に対し，ローブ幅  $N$  の Lanczos フィルタを垂直および水平方向に施す．

**手順 4** フィルタリングにより得られた出力画像を，表示方式に応じて，出力デバイスの中央もしくは中心点のオフセット値  $Offset(x, y)$  に応じて配置し，最終的な出力画像として表示する．

#### 3.2 画素アスペクト比変換

本節では，実装システムで行う画素アスペクト比変換について述べる．現在利用されている様々な映像表示デバイスをデジタルシネマと比較すると，空間解像度だけではなく画素アスペクト比も異なる場合がある．本論文では，画面アスペクト比を  $DAR$  (Display Aspect Ratio)，画素アスペクト比を  $PAR$  (Pixel Aspect Ratio) と表記する．

$DAR$  および  $PAR$  は，垂直方向を基準とした場合の水平方向の大きさを表している．すなわち，デジタルシネマのファイルフォーマットである DCDM (Digital Cinema Distribution Master) の 4K は， $DAR = 1.896 (= 4096/2160)$  と  $PAR = 1$  であり，NTSC 方式の SDTV は， $DAR = 1.33 (4 : 3)$  と  $PAR = 0.89 (8 : 9)$  である．

表 3 現行出力デバイスの DAR および PAR の値  
Table 3 Parameter values of DAR and PAR

	Horizontal[pe]l	Vertical[pe]l	DAR	PAR
HDTV	1920	1080	1.778	1.000
SDTV(NTSC)	720	480	1.333	0.889
SDTV(PAL)	720	576	1.333	1.067
VGA	640	480	1.333	1.000
QVGA	320	240	1.333	1.000
CIF	352	288	1.333	1.067
QCIF	176	144	1.333	1.067

表 3 に、現行の出力デバイスに対する  $DAR$  と  $PAR$  の値を示す。

解像度変換の前後において  $DAR$  と  $PAR$  が異なる場合、必ずしも出力デバイスの全画素を利用して表示できるとは限らない。そのため、解像度変換後の有効画素領域の表示方法が複数存在している。本システムでは、レターボックス方式とサイドカット方式を用いる。レターボックス方式は、映画などのコンテンツ、サイドカット方式は、スポーツなどのコンテンツに対して適しているとされる<sup>11)</sup>。

レターボックス方式では、入力画像の水平解像度を最大限に利用し、画面の上下には黒帯を入れる。したがって、解像度変換後の有効垂直画素数を、入力画像の  $DAR_s$ 、出力画像の  $PAR_d$ 、出力画像の水平解像度  $Dst(H)$  および垂直解像度  $Dst(V)$  を用いて式 (17) で表す。

$$Dst(V) = (Dst(H)/DAR_s) \times PAR_d \quad (17)$$

また、サイドカット方式では、入力画像の垂直解像度を最大限に利用するため、解像度変換後の画像は、両側の画素を切り取る必要がある。画素の切り取りを行う前の有効水平画素数を、式 (18) で表し、画像の片側で切り取る画素数  $Cut(H)$  を、表示デバイスの水平解像度  $Display(H)$  を用いて式 (19) で表す。

$$Dst(H) = (Dst(V) \times DAR_s) / PAR_d \quad (18)$$

$$Cut(H) = (Dst(H) - Display(H)) / 2 \quad (19)$$

#### 4. 評価実験

提案手法の有効性を示すために、画質評価実験を行う。本実験では、空間スケーラビリティを用いず、入力画像に直接 Lanczos フィルタを施す解像度変換法を従来手法とし、提案手法と比較する。

##### 4.1 評価方法と参照画像の定義

解像度変換により得られる縮小画像の品質を定量的に評価するためには、比較となる同じ縮小解像度の参照画像を定義する必要がある。本論文では、DFT (Discrete Fourier Transform) による解像度変換で得られる画像を、定量評価のための参照画像として定義する。DFT で得られる周波数成分のうち、解像度変換後に表現するこ

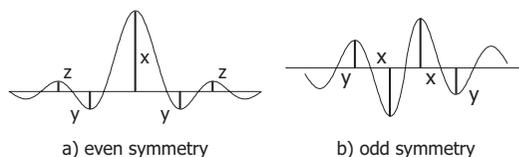


図 7 PPR の定義

Fig. 7 Definition of the PPR

とのできない高周波成分を除去することで、ダウンサンプリングによるエイリアシングノイズを防ぐことができ、変換後に必要な信号成分を維持することができる。

一方、DFT による解像度変換法では、離散周期信号の結合部分やステップ関数に近いエッジ周辺においてリングングが顕著に見られる場合がある。これは、離散周期信号の不連続性の影響によるもので、視覚的には理想的とは言い難い。そこで、リングングを定量評価するために PPR (Peak-to-peak Ratio) を用いる<sup>12)</sup>。PPR は、式 (20) で定義される値であり、値が大きいくほどリングングが抑制されることが知られている。通常、PPR が示す値の範囲は (1, 2] である。

提案手法では、システム全体のインパルス応答を偶対称の FIR フィルタで表現できるため、式 (20) の *even symmetry* を利用して PPR を算出する。式 (20) 中の  $x, y, z$  は、図 7 に示した距離に対応する。

$$PPR = \begin{cases} \frac{2(x+y)}{(x+y)+(y+z)}, & \text{even symmetry} \\ \frac{2x}{x+y}, & \text{odd symmetry} \end{cases} \quad (20)$$

##### 4.2 実験条件

実験には、*bride*, *harbor*, *bottles*, *wool* の 4 種類の標準画像<sup>13)</sup>を、 $4096 \times 2160$  画素に切り取って用いる。出力画像は、レターボックス方式とサイドカット方式の HDTV, VGA, QVGA, SDTV (NTSC), SDTV (PAL), CIF, QCIF の計 14 種類とする。解像度変換後に必要とされる有効画素領域の解像度は、式 (17) および式 (18) を用いて求める。ただし、DFT による解像度変換の出力結果と PSNR を正しく比較するために、解像度変換後の有効画素領域のサイズは偶数値に四捨五入して算出する。また、ビット深度は 8[bit/pe]l とし、式 (16) の  $\beta$  の値は、ゾーンプレートに対する実験結果である表 2 の値を用いる。入力画像の符号化には、JPEG 2000 Part1 で定義されている 9/7-Daubechies フィルタを利用する。

##### 4.3 実験結果

###### 4.3.1 PSNR による比較

表 4, 5 に、入力画像が *harbor* である場合の PSNR の測定結果を示す。表中  $L = 0$  の行で示された値が従来手法、 $L > 0$  で示された行が提案手法の PSNR である。2.1 節で示したように、低域画像とデシメーションフィル

表 4 画質比較 (harbor, Letter Box)  
Table 4 PSNR(harbor, Letter Box)

	HDTV	VGA	QVGA	NTSC	PAL	CIF	QCIF
L=0	42.28	35.62	33.81	35.65	35.98	34.14	32.45
L=1	<b>41.18</b>	36.96	35.04	36.95	<b>37.32</b>	35.44	33.56
L=2		<b>37.86</b>	<b>36.26</b>	<b>37.66</b>	37.76	<b>36.74</b>	34.66
L=3			36.11			36.06	<b>35.01</b>
L=4							34.24

[dB]

表 5 画質比較 (harbor, Side Cut)  
Table 5 PSNR(harbor, Side Cut)

	HDTV	VGA	QVGA	NTSC	PAL	CIF	QCIF
L=0	<b>43.07</b>	36.82	34.67	37.07	37.60	35.14	33.11
L=1	41.24	<b>38.08</b>	35.98	<b>37.54</b>	<b>38.82</b>	35.85	34.30
L=2		36.31	<b>37.00</b>	35.22		<b>37.04</b>	35.22
L=3			34.14				<b>35.51</b>

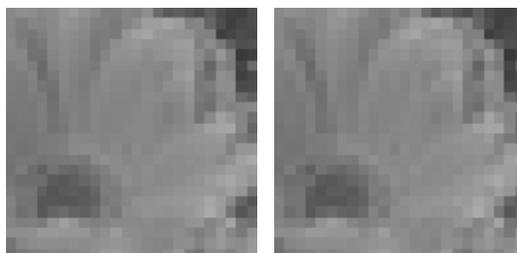
[dB]

タの組み合わせが複数存在するため、すべての場合を示している。表中の太字で示した値が、実際の提案手法における低域画像とデシメーションフィルタの組み合わせの結果である。実験結果より、HDTV を除く解像度において、従来手法よりも 1~2[dB] の画質が改善していることがわかる。これは、演算量を考慮した上でよりタップ数の多い Lanczos フィルタを利用したためであると考えられる。一方で、選択可能な最大分割数  $L_{max}$  が 1 である場合には、必ずしも低域画像との組み合わせが良いとは限らないと言える。

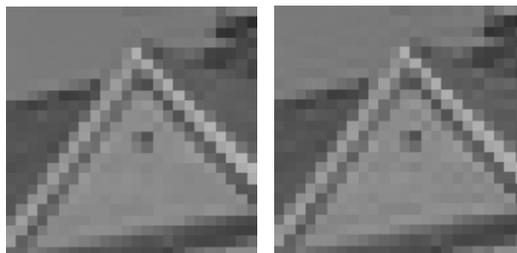
#### 4.3.2 出力画像の比較

提案手法の有効性を確認するために、実際の出力画像の比較を行った。まず、入力画像がデジタルシネマサイズの DCDM (4096 × 2160 画素)、出力画像がレターボックス方式の QVGA である場合について、出力結果を図 8 に示す。図 8 は、従来手法との差を明確にするために、QVGA の出力画像の一部 (24 × 24 画素) を切り出し、拡大している。左側が従来手法、右側が提案手法による出力結果である。図 8-(a) および図 8-(c) より、提案手法を用いることで従来手法では平滑化されてしまう精細部分を保存していることがわかる。また、図 8-(d) では、斜めのエッジが強調され、出力画像の全体としては引き締まった印象を受ける。一方で、図 8-(b) のようにエッジが平坦な領域に面している場合は、特に水平、垂直方向へのリングングが観測されるが、従来手法との比較では、全体的にシャープな出力画像が得られた。

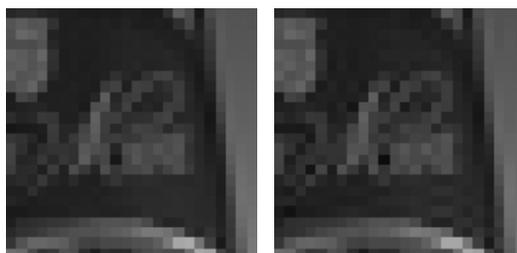
次に、入力画像がデジタルシネマサイズの DCDM (4096 × 2160 画素)、出力画像がレターボックス方式の PAL である場合について、出力結果を図 9 に示す。図 9 は、PAL の出力画像の一部 (32 × 32 画素) を切り出し、



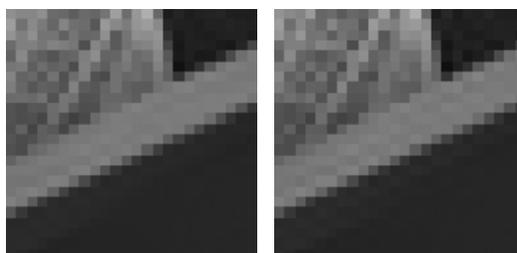
conventional (L=0)      proposed (L=2)  
(a) bride



conventional (L=0)      proposed (L=2)  
(b) harbor



conventional (L=0)      proposed (L=2)  
(c) bottles



conventional (L=0)      proposed (L=2)  
(d) wool

図 8 従来手法と提案手法の比較 (QVGA)  
Fig. 8 Experiment results on natural images (QVGA)

拡大している。DCDM から PAL への変換の場合、従来手法との明確な差がわかるのは、図 9 に示すようなエッジ周辺の変化であった。提案手法を用いることで、エッジが強調されていることがわかる。また、主観評価では、

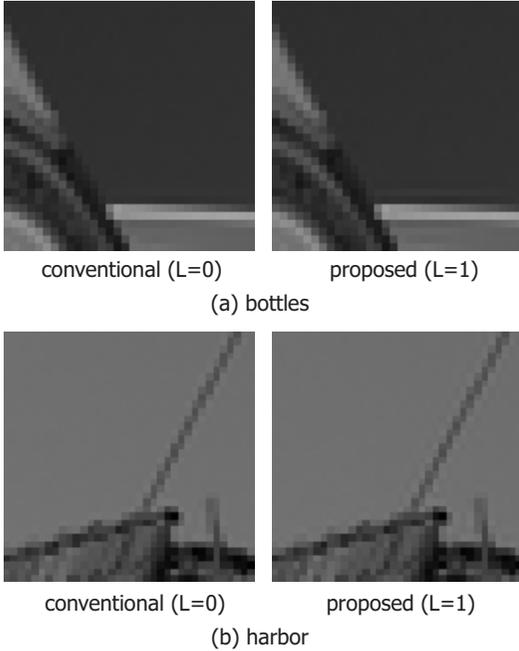


図9 従来手法と提案手法の比較 (PAL)

Fig. 9 Experiment results on natural images (PAL)

目立ったリングングは観測されず、提案手法により全体的にシャープな出力画像が得られることを確認した。

最後に、提案手法により得られる画像の性質を定量的に評価するために、PPRの値を比較する。表6, 7に、従来手法と提案手法のPPRの値を示す。4.3.1節の実験結果においてPSNRの値が改善しなかったHDTVの場合を除き、提案手法のPPRは従来手法よりも低いことが確認できる。これは、提案手法が、従来手法と比較してリングングが抑制されにくいことを表しており、主観評価において出力画像に若干のリングングが観察されたことに一致する。ここで、提案システムで利用したLanczosフィルタのローブ幅 $N$ とPPRとの関係を図10に示す。図10は、入力画像がharborで、出力画像がレターボックス方式である場合のプロット結果であり、実線で結ばれたプロット点は同一の出力解像度である。図10より、HDTVを除き、 $N$ とPPRとの間に、 $N$ が増加するにつれ、PPRの値が減少するという相関性が見られる。したがって、PPRの値を利用することで、ある程度提案システムの出力画像の特性を調節することができると考えられる。

#### 4.3.3 提案システムの正答率に対する評価

本節では、提案手法が、複数ある低域画像とデシメーションフィルタの組み合わせから、最高品質を得る組み合わせを選択する正答率を評価する。4.3.1節の実験結

表6 PPR (harbor, Letter Box)  
Table 6 PPR (harbor, Letter Box)

	HDTV	VGA	QVGA	NTSC	PAL	CIF	QCIF
従来	1.774	1.732	1.732	1.732	1.732	1.731	1.731
提案	1.819	1.537	1.585	1.537	1.631	1.577	1.565

表7 PPR (harbor, Side Cut)  
Table 7 PPR (harbor, Side Cut)

	HDTV	VGA	QVGA	NTSC	PAL	CIF	QCIF
従来	1.787	1.736	1.732	1.736	1.746	1.732	1.731
提案	1.787	1.644	1.596	1.730	1.633	1.563	1.547

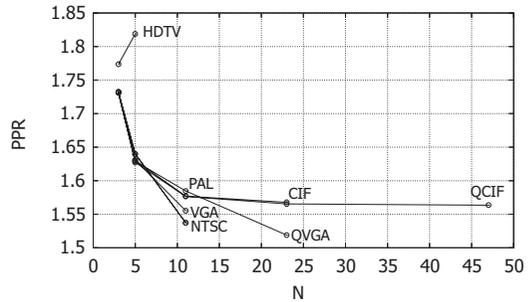


図10 PPRとNの関係

Fig. 10 Shift of PPR due to N

果より、正答率は実験で使用した画像の種類に依存せず85.7%であった。より正確に最高品質を与える組み合わせを推定するためには、入力画像の性質に応じた $\beta$ の利用や、システム全体の周波数特性を詳細に解析する必要がある。表4に示した、正しい選択が行われなかったレターボックス方式のSDTV(PAL)への変換では、判定誤りによるPSNRの差がほぼ0.5[dB]程度に抑えられている。同時に、最高品質には至らないが、従来手法よりも高いPSNRの値を示している。また、判定誤りで出力される縮小画像と正しい結果との差分画像を生成し、主観評価を行った結果、両者が知覚できないレベルの差であることを確認した。加えて、図9-(a)および(b)の提案手法の出力画像は、判定誤りの出力結果であるが十分な画質を得ている。したがって、判定誤りで選択される画像を用いても、実用上の問題にはならないと言える。

以上より、提案する空間スケーラビリティを用いた解像度変換手法が有効な手法であることがわかる。

## 5. むすび

本論文では、JPEG 2000の空間スケーラビリティから得られる低域画像とLanczosフィルタによる縮小処理により、有理数比の解像度変換を行う手法について述べた。提案手法は、デシメーションフィルタの演算量を定

式化することで、従来手法よりも低い演算量での処理を保証した。提案手法を用いることで、原画像に直接デシメーションフィルタをかける場合と比較して、1~2[dB]程度の画質改善を実現した。

謝辞

本研究は、情報通信研究機構(NICT)「通信ネットワーク利用放送技術の研究開発」の研究課題による。

参考文献

- 1) 永吉功, 笠井裕之, 花村剛, 富永英義, “再量子化/再符号化処理共有によるマルチレート出力ビデオトランスコーダ,” 電子情報通信学会論文誌, vol.J86-B, no.12, pp.2511-2521, December 2003.
- 2) J.W. Woods, and G. Lilenfield, “A Resolution and Frame-Rate Scalable Subband/Wavelet Video Coder,” IEEE Trans. on Circuits and Systems for Video Technology, vol.11, no.9, pp.1035-1044, 2001.
- 3) Digital Cinema Initiatives, LLC, “Digital Cinema System Specification v5.0,” March 2005.
- 4) ISO/IEC JTC1/SC29 WG1, “ISO/IEC 15444-1 JPEG 2000 Part1: Core coding system,” July 2002.
- 5) 富永英義, 川島正久, 沼田文彦, “M バンド拡張 QMF バンクを用いた動画の帯域分割符号化,” 電子情報通信学会論文誌, vol.J74-B-I, no.12, pp.1065-1073, December 1991.
- 6) 高橋靖正, 田口亮, “高周波数成分推定を伴う任意倍率可能な画像拡大法,” 電子情報通信学会論文誌, vol.J84-A, no.9, pp.1192-1201, September 2001.
- 7) 仲地孝之, 澤邊知子, 鈴木純司, 藤井哲郎, “フィルタバンクを用いた非オクターブ型スケーラブル符号化法の検討,” 信学技報, pp.23-38, April 2004.
- 8) K. Turkowski, “Filters for Common Resampling Tasks,” Technical report, Apple Computer, October 1990.
- 9) R.E. Crochiere, and L.R. Rabiner, Multirate Digital Signal Processing, PRENTICE-HALL, 1983.

- 10) 尾知博, シミュレーションで学ぶデジタル信号処理, CQ 出版社, 2001.
- 11) 神原康伸, “アップ/ダウンコンバートの基礎知識,” ビデオ, pp.52-64, June 2003.
- 12) E. A. B. da Silva, and M. Ghanbari, “On the Performance of Linear Phase Wavelet Transforms in Low Bit-Rate Image Coding,” IEEE Trans. on Image Processing, vol.5, no.5, pp.689-704, May 1996.
- 13) 画像電子学会, “高精細 XYZ・CIELAB・RGB 標準画像 (SHIPP),” April 2000.

(2005年5月9日受付)

(2005年7月20日再受付)

石川 孝 明



平 15 早大・理工・電子通信卒。平 17 同大大学院修士課程了。現在、同後期博士課程在学中。主として、画像符号化に関する研究に従事。電子情報通信学会、画像電子学会各学生会員。

渡 辺 裕 (正会員)



昭 55 北大・工・電子卒。昭 60 同大学院博士課程修了。同年 NTT 入社。NTT 研究所において画像の高効率符号化に関する研究及び MPEG 標準化に従事。平 12 早稲田大学大学院国際情報通信研究科教授。マルチメディアの符号化及び配信に関する研究に興味を持つ。工博。IEEE, 電子情報通信学会, 情報処理学会, 映像情報メディア学会各会員。